JOURNAL OF ALGEBRA **84**, 42–61 (1983)

Embedding Rings in Completed Graded Rings
3. Algebras over General k

G. M. BERGMAN,* D. J. BRITTEN,† AND F. W. LEMIRE†

*University of California, Berkeley, California 94720, and
†University of Windsor, Windsor, Ontario N9B 3P4, Canada

Communicated by I. N. Herstein

Received February 1, 1982

It is shown that any associative algebra R over a commutative ring k such that $R^{n+1} = \{0\}$ can be embedded in a graded k-algebra $\bigoplus H_i$, such that $H_i \neq \{0\}$ only for $1 \leqslant i \leqslant 2^{n-1}$. By the results of the first paper in this series, R can therefore be embedded in $(2^{n-1}+1) \times (2^{n-1}+1)$ strictly upper triangular matrices over a commutative k-algebra. For k any integral domain *not* a field, this is in fact a best result, in contrast with the case of k a field, where one can replace "2^{n-1}" by "n" (see Part 2 of this series). More generally, if R is a nonunital k-algebra with a strictly positive-integer-valued filtration function, then R can be mapped into a completed $\mathbb{Z}$ graded k algebra S by a homomorphism f such that $v(a) \leqslant v(f(a)) \leqslant 2^{v(a)-1}$, where v denotes the given filtration function on R and the grading-filtration on S. If all values of the filtration function on R are powers of 2, one can even get $v(f(a)) = v(a)$.

INTRODUCTION

The problem of embedding nilpotent rings in rings of strictly upper triangular matrices $T_m(C)$ over commutative rings C has a long history. Some of the first results on this problem were given by Engel for the case of Lie algebras [5, Sections 3.1–3.2], and by Shaw, who showed that any n-dimensional nilpotent algebra over a field k is embeddable in $T_{n+1}(k)$ [6, Theorem 2.1.2, p. 17]. Further results in this direction are obtained in the preceding papers [1] and [2] of this series; see also references to related work of Anan'in and L'vov in [2, (5.5)].

In this paper we show that every nilpotent associative algebra over a commutative ring k is embeddable in a strictly upper triangular matrix ring over some commutative k-algebra, as well as establishing the other embedding theorems indicated in the abstract. As in [1] and [2], our results are proved by examining the universal object for this problem. But where the approach of [1] was to chop away at the universal object till one was left with something one could describe, and that of [2] was to assume k a field

0021-8693/83 $3.00

and slice up the universal object using module splittings, we shall here handle it with kid gloves—albeit these gloves will be attached to a Rube Goldberg machine. This "machine" produces certain magical systems of integer coefficients, which we use in defining linear maps on the universal object, which demonstrate the desired properties.

We begin with examples showing that one cannot do better than the results stated in the abstract. In Section 2–3 we prove the embedding theorem, limiting ourselves for conceptual simplicity to the result on nilpotent algebras. But in Section 5 we show that the systems of integer coefficients constructed for the nilpotent case can be applied to get the more general statement about filtered algebras.

1. Conventions; Counterexamples

Throughout this paper, k will be a commutative ring with 1; k-algebras will be associative but need not have 1.

We shall assume the definitions regarding filtered and graded algebras (filtered or graded by $\mathbb{Z}$) of [2, Sections 1–3]. Concerning graded algebras, the main point to observe is notational: a graded algebra is taken to be its system of homogeneous components, $H = (H_i)$ $(i \in \mathbb{Z})$, while we write $H^{\oplus}$ for the total algebra $\bigoplus H_i$, and $H^{\widehat{\oplus}}$ for its completion.

The following Lemma shows a way in which the filtration function on an algebra $H^{\widehat{\oplus}}$ is more restricted than that of a general filtered k-algebra.

Lemma 1. *Suppose H is a $\mathbb{Z}$-graded k-algebra, and $x \in H^{\widehat{\oplus}}$, $t \in k$ are elements such that*

$$x^2 = tx, \tag{1.1}$$

$$v(x) \geqslant 1. \tag{1.2}$$

Then for $n \geqslant 1$,

$$v(x^n) \geqslant 2^{n-1}. \tag{1.3}$$

Proof. For $n = 1$, (1.3) is just (1.2). Assume (1.3) for n. By (1.1) we can rewrite (1.3) as

$$v(t^{n-1}x) \geqslant 2^{n-1}. \tag{1.4}$$

Hence letting y denote the sum of the homogeneous components of x of degrees $<2^{n-1}$ we have

$$x = y + z, \tag{1.5}$$

$$v(z) \geqslant 2^{n-1}, \tag{1.6}$$

$$t^{n-1}y = 0. \tag{1.7}$$

We now see that $x^{n+1} = t^{n-1}x^2 = t^{n-1}(y+z)^2 = t^{n-1}z^2$ by (1.7), so $v(x^{n+1}) = v(t^{n-1}z^2) \geqslant 2v(z) \geqslant 2^n$, proving (1.3) for $n+1$. ∎

From this we can deduce that the results of [2] fall far short of holding over base-rings that are not fields:

COROLLARY 2. *Suppose k is an integral domain, and $t \in k$ a nonzero nonunit.*

(i) *Let n be any positive integer and $R = tk/t^{n+1}k$ (quotient of an ideal of k by a subideal, considered as a nonunital k-algebra). Then $R^{n+1} = \{0\}$, but if R can be embedded in a k-algebra $H^{\oplus} = H_1 \oplus \cdots \oplus H_m$ (i.e., all other components zero) then m must be $\geqslant 2^{n-1}$.*

(ii) *Suppose $\bigcap_i t^i k = \{0\}$, and let R denote the ideal $tk \subseteq k$, considered as a nonunital k-algebra, and filtered by letting $R_{(i)} = t^i k$ $(i \geqslant 1)$. Then for every positive integer n, $v(t^n) = n$, but under any homomorphism f of filtered k-algebras from R into an algebra $H^{\hat{\oplus}}$, $v(f(t^n)) \geqslant 2^{n-1}$.*

Proof. To get the final assertion in each case let x denote the image of t in $H^{\hat{\oplus}}$ ($= H^{\oplus}$ in case (i)), and apply Lemma 1. ∎

The related example in Corollary 4 below shows that if instead of filtered algebras with strictly *positive* filtration functions, we allow *nonnegative* filtration functions, then no results of the sort we want to prove hold—again in contrast to the case of k a field, where we could allow arbitrary $\mathbb{Z}$-valued filtrations [2].

LEMMA 3. *Suppose H is a graded algebra, and $e \in H^{\hat{\oplus}}$ an idempotent element. Then for all $t \in k$, $v(te)$ is either $\leqslant 0$ or $+\infty$.*

Proof. Suppose $v(te) \neq +\infty$. Let us write $e = y + z$, where y is the sum of the homogeneous components of e of degrees $< v(te)$. Thus $ty = 0$, and $v(z) \geqslant v(te)$. Hence $te = te^2 = t(y+z)^2 = tz^2$, so $v(te) \geqslant 2v(z) \geqslant 2v(te)$. So $0 \geqslant v(te)$. ∎

COROLLARY 4. *Suppose k, t are as in Corollary* 2(ii), *and now let $R = k$, again filtered by $R_{(i)} = t^i k$ $(i \geqslant 0)$. Then in R, $v(t) = 1$, but any homomorphism f of filtered k-algebras from R into an algebra $H^{\hat{\oplus}}$ annihilates t.*

Proof. In $H^{\hat{\oplus}}$, $f(1)$ is idempotent, and $v(tf(1)) = v(f(t)) \geqslant v(t) = 1$. Hence $v(f(t)) = v(t(f(1))) = +\infty$; i.e., $f(t) = 0$. ∎

(1.8) We have not attempted to give the above examples in the most general possible form. For instance, in Lemma 1 we might have replaced x by any additive subgroup X satisfying $X^2 \subseteq tX$. We can get an example like that of Corollary 2(i) over any commutative ring k *unless* the quotient k/N of k by its nil radical N is von Neumann regular and $N^{n-1} = \{0\}$. In (ii) we could (for starters) have deleted the hypothesis $\bigcap_i t^i k = \{0\}$ and instead taken $R = tk/\bigcap_i t^i k$. In Corollary 4, as long as k itself was not von Neumann regular we could have found a t such that $t \notin t^2 k$, and let $R = k/t^2 k$, with the (t)-adic filtration, here $\{0, 1, +\infty\}$-valued, and the same conclusion would have held.

2. Setting up the Problem

Given a commutative ring k, and given positive integers n and m, we wish to know whether the following holds:

> Every associative k-algebra R satisfying $R^{n+1} = \{0\}$ can be embedded in a k-algebra $H^{\oplus}$, where H is a $\mathbb{Z}$-graded k-algebra zero in all degrees except $1, \ldots, m$. (2.1)

Now for any k-algebra R, the universal algebra-homomorphism $\varphi: R \to U^{\oplus}$ where U is a graded k-algebra zero in all degrees except $1, \ldots, m$ can be constructed as follows. Let T be the graded k-algebra freely generated by m k-modules isomorphic to R, one in each of degrees $1, \ldots, m$, which we shall write

$$\theta_1(R) \subseteq T_1, \ldots, \theta_m(R) \subseteq T_m, \qquad (\theta_i : R \to T_i). \tag{2.2}$$

This T is a "graded tensor algebra"; its component T_i $(i \geqslant 1)$ is the direct sum of all tensor product modules

$$\theta_{j(1)}(R) \otimes_k \cdots \otimes_k \theta_{j(r)}(R) \qquad \text{with} \qquad j(1) + \cdots + j(r) = i \quad (r \geqslant 1, j(p) \leqslant m). \tag{2.3}$$

T is characterized by the property of having a universal *k-module homomorphism* $\theta: R \to T^{\oplus}$ whose image lies in degrees $1, \ldots, m$:

$$\theta(x) = \theta_1(x) + \cdots + \theta_m(x) \in T^{\oplus}. \tag{2.4}$$

To get our U, we "truncate" T by replacing all T_i such that $i > m$ with $\{0\}$, and then impose the relations needed to turn (2.4) into an algebra homomorphism. Given that it is already a module homomorphism, this means that we must divide out by the ideal generated by the elements

$$\theta_i(xy) - \sum_{\mu < i} \theta_\mu(x)\,\theta_{i-\mu}(y). \tag{2.5}$$

We now define $\varphi: R \to U^{\oplus}$ to be the composition $R \to^{\theta} T^{\oplus} \to U^{\oplus}$ of θ with this quotient-map. This φ will be the asserted universal homomorphism.

We note that the highest homogeneous component, U_m, will be the factor-module of T_m by the submodule spanned by all elements

$$\theta_{i(1)}(z_1) \otimes \theta_{i(2)}(z_2) \otimes \cdots \otimes \theta_{i(p)}(xy) \otimes \cdots \otimes \theta_{i(q)}(z_q)$$
$$- \sum_{\mu} \theta_{i(1)}(z_1) \otimes \cdots \otimes \theta_\mu(x) \otimes \theta_{i(p)-\mu}(y) \otimes \cdots \otimes \theta_{i(q)}(z_q), \tag{2.6}$$

$(1 \leqslant p \leqslant q;\ i(1) + \cdots + i(q) = m;\ x, y$ and all z's in R. Cf. (2.3), (2.5)).

Now for each string of positive integers $j(1),\ldots,j(r)$ summing to m, let

$$\pi_{j(1),\ldots,j(r)} : T_m \to R$$

denote the k-linear map which on the summand $\theta_{j(1)}(R) \otimes \cdots \otimes \theta_{j(r)}(R)$ is the map induced by the multiplication of R, carrying $\theta_{j(1)}(z_1) \otimes \cdots \otimes \theta_{j(r)}(z_r)$ to $z_1 \cdots z_r \in R$, while it is the zero map on all other summands $\theta_{j'(1)}(R) \otimes \cdots \otimes \theta_{j'(r')}(R)$. Suppose we can find a k-linear combination

$$\alpha = \sum a_{j(1),\ldots,j(r)} \pi_{j(1),\ldots,j(r)} \qquad (a_{j(1),\ldots,j(r)} \in k) \tag{2.7}$$

of these maps, which annihilates all of the relators (2.6). Then α will induce a k-linear map $\bar{\alpha}$: $U_m \to R$:

$$\begin{tikzcd} & T_m \arrow[dr, "\alpha"] \arrow[dd] & \\ R \arrow[ur, "\theta_m"] \arrow[dr, "\varphi_m"'] & & R. \\ & U_m \arrow[ur, "\bar{\alpha}"'] & \end{tikzcd}$$

Suppose that α can further be chosen so that

$$\alpha \circ \theta_m = id_R. \tag{2.8}$$

This gives $\bar{\alpha} \circ \varphi_m = id_R$, implying that φ_m is one-to-one, hence $\varphi: R \to U^{\oplus}$ will be so as well, and (2.1) will be established.

Clearly, α will annihilate the particular relator shown in (2.6) if

$$a_{i(1),\dots,i(p),\dots,i(q)} = \sum_{\mu} a_{i(1),\dots,\mu,i(p)-\mu,\dots,i(q)}. \tag{2.9}$$

Note further that if we limit our attention to rings R satisfying

$$R^{n+1} = \{0\}, \tag{2.10}$$

then all $\pi_{j(1),\dots,j(r)}$ with $r > n$ will be zero, so we may as well restrict the sum (2.7) to summands with $r \leqslant n$. Moreover, if $q > n-1$, (2.10) implies that the relator (2.6) is annihilated by all the $\pi_{j(1),\dots,j(r)}$, so for α to annihilate all our relators we only need require (2.9) in the case $q \leqslant n-1$. Note, finally, that (2.8) will hold if $a_m = 1$ (where m is, of course, the unique length-1 string summing to m). Summarizing, we have

LEMMA 5. *Let $J_{n,m}$ denote the set of all sequences of positive integers $(j(1),\dots, j(r))$ of lengths $r \leqslant n$, satisfying $j(1) + \cdots + j(r) = m$.*

Then a sufficient condition for the embedding property (2.1) *to hold is that there exist a $J_{n,m}$-tuple of elements $a_{j(1),\dots,j(r)} \in k$, such that for all $(i(1),\dots, i(q)) \in J_{n-1,m}$ and all p with $1 \leqslant p \leqslant q$,*

$$a_{i(1),\dots,i(p),\dots,i(q)} = \sum_{1 \leqslant \mu < i(p)} a_{i(1),\dots,\mu,i(p)-\mu,\dots,i(q)} \tag{2.11}$$

and such that

$$a_m = 1. \tag{2.12}$$ ∎

Note that if we have such a system of elements $a_{j(1),\dots,j(r)}$ for a given ring k, we automatically have one for any k' into which k can be mapped homomorphically. Since any nonzero commutative ring k can be mapped homomorphically *into* an integral domain which is not a field (e.g., a polynomial ring over a residue field of k), Corollary 2(ii) gives us for any nonzero k the lower bound $m \geqslant 2^{n-1}$ for (2.11), (2.12) to have solutions in k.

Let us get some familiarity with these systems by looking at examples for small n with $m = 2^{n-1}$:

(2.13) $n = 2$, $m = 2$. Then we need elements $a_2, a_{11} \in k$ satisfying $a_2 = a_{11}$ and $a_2 = 1$. These conditions obviously determine a_2 and a_{11}. The corresponding map α is given by $\alpha = \pi_2 + \pi_{11}$.

(2.14) $n = 3$, $m = 4$. Then we need elements

$$a_4, a_{13}, a_{22}, a_{31}, a_{112}, a_{121}, a_{211}$$

satisfying the following equations. (When two equations are given on a line, they are instances of (2.11) differing only in the value of p.)

$$a_4 = a_{13} + a_{22} + a_{31},$$

$$a_{13} = 0, \qquad a_{13} = a_{112} + a_{121},$$

$$a_{22} = a_{112}, \qquad a_{22} = a_{211},$$

$$a_{31} = a_{121} + a_{211}, \qquad a_{31} = 0.$$

$$a_4 = 1.$$

We immediately get the unique solution

$$a_4 = a_{22} = a_{112} = a_{211} = 1, \qquad a_{13} = a_{31} = 0, \qquad a_{121} = -1,$$

i.e.,

$$\alpha = \pi_4 + \pi_{22} + \pi_{112} - \pi_{121} + \pi_{211}.$$

Note how the equations $a_{13} = 0$ and $a_{31} = 0$ arose: as cases of (2.11) with $i(p) = 1$, making the right-hand side of that equation the empty sum. We see in the same way that, generally, a term $a_{i(1),...,i(q)}$ to which (2.11) applies, i.e., one for which the number of subscripts, q, is $\leqslant n-1$, will be zero unless all of these subscripts are $\geqslant 2$. Consider next a term $a_{i(1),...,i(q)}$ with $q = n-2$. If any of its subscripts $i(p)$ is <4, then when we expand by (2.11) with respect to this subscript, all terms on the right have $n-1$ subscripts, at least one of which (one of μ, $i(p) - \mu$) is <2, so this term is also 0. By induction we get

> If the $J_{n,m}$-tuple of elements $a_{j(1),...,j(r)} \in k$ satisfies (2.11), then each of its *nonzero* terms $a_{j(1),...,j(r)}$ must have $j(1),..., j(r) \geqslant 2^{n-r}$. (2.15)

Observe that taking $r = 1$, this implies

> If the system of elements $a_{j(1),...,j(r)}$ satisfies (2.11) and (2.12), and $k \neq \{0\}$, then $m \geqslant 2^{n-1}$, (2.16)

which is just what we deduced above from the counterexamples of Section 1.

The next case we shall only sketch. It is notable as the first case in which the coefficients are not unique.

(2.17) $n = 4$, $m = 8$ (sketch). By (2.15) the only a's that can be nonzero are

> a_8, a_{44}, a_{422}, a_{242}, a_{224}, a_{332}, a_{323}, a_{233}, and *the* 35 *terms* with subscripts of length $q = 4$.

Of the latter 35 terms, a_{1115}, a_{1151}, a_{1511}, a_{5111} can be seen to be zero by computations such as: $a_{1115} = a_{215} = 0$ by (2.15). The solution of (2.11) and (2.12) for the remaining terms is tedious, but is made somewhat easier by the fact that many of these equations reduce to saying that two terms have sum or difference 0. For example, the equation $a_{35} = a_{314} + a_{323} + a_{332} + a_{341}$ becomes, on applying (2.15), $0 = a_{323} + a_{332}$.

One obtains, in fact, a solution with five degrees of freedom. The solution involving the fewest nonzero terms, expressed as a formula for α, is

$$\begin{aligned}\alpha = \pi_8 + \pi_{44} + \pi_{422} - \pi_{242} + \pi_{224}\\ + \pi_{4211} + \pi_{4112} - \pi_{2411} - \pi_{1142} + \pi_{2114} + \pi_{1124}\\ - \pi_{4121} - \pi_{1214}\\ + \pi_{2222} + \pi_{2321} - \pi_{2312} + \pi_{1232} - \pi_{2132}.\end{aligned}$$

To exhibit one of the degrees of freedom, we note that any multiple of the functional $\pi_{2222} - \pi_{2213} - \pi_{2132} + \pi_{2123} - \pi_{1322} + \pi_{1313} + \pi_{1232} - \pi_{1223}$, when added to a solution, gives another solution.

3. The General Solution

We shall now show how to obtain all solutions to (2.11) over any k, for any n and m. In particular, we shall see that whenever $m \geqslant 2^{n-1}$ there exist solutions satisfying (2.12).

Our construction will work progressively from elements $a_{j(1),\ldots,j(r)}$ with lower subscript-length r to those with higher r, and for a given r will work inductively by "pushing to the right" occurrences of the smallest subscript allowed by (2.15), 2^{n-r}. Let us set up an index on which to perform induction. For $(j(1),\ldots, j(r)) \in J_{n,m}$, we define $h(j(1),\ldots, j(r))$ to be -1 if any of the indices is $<2^{n-r}$; otherwise we define it to be the sum, over all subscripts $j(p)$ that are exactly 2^{n-r}, of their distance from the right-hand end of the string, $r-p$:

$$h = h(j(1),\ldots, j(r)) = \begin{cases} -1 & \text{if any } j(p) < 2^{n-r}, \\ \displaystyle\sum_{j(p)=2^{n-r}} r-p & \text{otherwise.} \end{cases} \tag{3.1}$$

So, for instance, $h(j(1),\ldots, j(r)) = 0$ if and only if all $j(p)$ are $\geqslant 2^{n-r}$, with strict inequality except possibly for the last term, $j(r)$.

Assume inductively that the elements $a_{j'(1),\ldots,j'(r')}$ have been defined for all $r' < r$, so as to satisfy all instances of (2.11) with $q < r-1$, and also the conclusion of (2.15). We assign the values $a_{j(1),\ldots,j(r)}$ as follows:

(3.2) If $h(j(1),\ldots, j(r)) = -1$, we set $a_{j(1),\ldots,j(r)} = 0$, as required by (2.15).

(3.3) If $h(j(1),\ldots, j(r)) = 0$, we choose $a_{j(1),\ldots,j(r)}$ arbitrarily!

If $h(j(1),\ldots, j(r)) > 0$, let us also assume inductively that $a_{j'(1),\ldots,j'(r)}$ has been defined for all subscripts of length r satisfying $h(j'(1),\ldots, j'(r)) < h(j(1),\ldots, j(r))$. Now the condition $h(j(1),\ldots, j(r)) > 0$ means in particular that for some $p < r$, $j(p) = 2^{n-r}$. Consider the case of (2.11) with $q = r-1$, a value of p such that $j(p) = 2^{n-r}$, and $(i(1),\ldots, i(q)) = (j(1),\ldots, j(p) + j(p+1),\ldots, j(r))$. Note that the $\mu = 2^{n-r}$ term in the right-hand sum of (2.11) is precisely the element we want to define. We claim also that all other terms of this sum have subscripts with lower values of h, and hence have already been defined.

Indeed, those with $\mu < 2^{n-r}$ or $\mu > i(p) - 2^{n-r}$ have $h = -1$, and so have been set to zero by (3.2). Those with $2^{n-r} < \mu < i(q) - 2^{n-r} = j(p+1)$ have lost the 2^{n-r} in the pth place of their subscript, and not gained one anywhere else, so they certainly have smaller h. The term with $\mu = i(p) - 2^{n-r}$ (if distinct from the term with $\mu = 2^{n-r}$, i.e., if $j(p+1) > 2^{n-r}$) has traded its subscript 2^{n-r} in the pth place for one in the $p+1$st place, thus decreasing h by 1.

The left-hand side of (2.11) has already been defined because the subscript there has length $r-1$.

Hence we shall use this instance of (2.11) to *define* $a_{j(1),\ldots,j(r)}$. Discarding terms of the sum which are 0 by (2.15), this means

(3.4) If $h(j(1),\ldots, j(r)) > 0$, with $j(p) = 2^{n-r}$, $p < r$, we define

$$a_{j(1),\ldots,j(r)} = a_{j(1),\ldots,2^{n-r}+j(p+1),\ldots,j(r)} - \sum_{2^{n-r}<\mu\leqslant j(p+1)} a_{j(1),\ldots,\mu,2^{n-r}+j(p+1)-\mu,\ldots,j(r)}.$$

The one difficulty is that among $j(1),\ldots, j(r-1)$ there may be more than one term equal to 2^{n-r}. So we must show that even if this happens, the resulting values given by (3.4) coincide.

Before undertaking this calculation, let us make a convention that will help tame typographical monstrosities such as (3.4). Let us agree to suppress all indices except those that will vary in a given calculation, and to denote the corresponding "a" term *by those indices, written in brackets*. The indices to be written will be called "distinguished indices," and will be specified in each case. For instance, below we shall rewrite (3.4) taking the pth and $p+1$st indices of the term on the left as distinguished. This means that any string of indices in brackets will denote the "a" term whose subscript consists of $j(1),\ldots,j(r)$ with the two terms $j(p)$, $j(p+1)$ replaced by the indicated string (possibly of different length). Let us for the remainder of this section use the abbreviation

$$2^{n-r}=c, \tag{3.5}$$

and in rewriting (3.4) let us also write $j(p+1)=d$. Then (3.4) becomes

$$[c,d]=[c+d]-\sum_{c<\mu\leqslant d}[\mu,c+d-\mu]. \tag{3.6}$$

When, as in case 1 below, the distinguished indices do not necessarily form a consecutive string, we will separate possibly nonconsecutive parts with semicolons.

We should keep in mind that a term whose subscript has length r must have all indices at least c if it is to be nonzero, while the corresponding condition on a term whose subscript is *shorter* by one index is that all indices be at least $2c$. One can deduce from this that in those cases where we are *not* making (3.6) a definition, i.e., where the hypothesis of (3.4) does not hold, (3.6) is nonetheless satisfied, as a consequence of our other defintions. Namely, since the term $[c,d]$ has a c in nonfinal position, the only way it could have been defined other than by (3.4) is by (3.2). In that case, there are two possibilities: $d<c$, or some other $j(p')<c$. In either case, (3.6) is easily seen to reduce to $0=0$. (In the first case, our observation on shortened subscripts is used.)

We are now ready to test the definition (3.4) for consistency when the string $j(1),\ldots,j(r)$ has more than one index equal to c. Say

$$j(p)=j(p')=c, \qquad p<p',$$

and assume inductively that

> Every "a"-term whose subscript has length $<r$, or length r and lower value of h, has already been shown to be consistently defined by (3.2)–(3.4). (3.7)

We consider two cases.

Case 1. $p'-p>1$. Then taking p, $p+1$, p', $p'+1$ as distinguished indices in $a_{j(1),\ldots,j(r)}$, and writing $j(p+1)=d$, $j(p'+1)=d'$, this element will be written as

$$[c,d;c,d'].$$

If we apply (3.6) at the first pair of indices, this becomes

$$[c+d; c, d'] - \sum_{c<\mu\leqslant d} [\mu, c+d-\mu; c, d'].$$

We note that the term preceding the sum equals zero, because its subscript has been shortened but it still has an index equal to $c < 2c$. In the summation every index has h-value less than the value of the original term, so by inductive hypothesis (3.7) we may apply (3.6) again, at the second pair of indices. This gives

$$-\sum_{c<\mu\leqslant d} [\mu, c+d-\mu; c+d'] + \sum_{c<\mu\leqslant d, c<\nu\leqslant d'} [\mu, c+d-\mu; \nu, c+d'-\nu]. \tag{3.8}$$

Now in the first sum in (3.8), we may drop all summands such that μ or $c+d-\mu$ is $<2c$. This leaves us with the range of summation $2c \leqslant \mu \leqslant d-c$, and hence a summation to which we can apply (3.4), equivalently (3.6), with $r-1$ for r and $2c$ for c. So (3.8) becomes

$$[c+d; c+d'] + \sum_{c<\mu\leqslant d, c<\nu\leqslant d'} [\mu, c+d-\mu; \nu, c+d'-\nu]. \tag{3.9}$$

We see that this is "symmetrical" with respect to the p and p'-locations. Hence the calculation starting with the p'-reduction gives the same result, as required.

Case 2. $p' = p+1$. Then we take $p, p', p'+1$ as distinguished indices, put $j(p'+1) = d$, and write $a_{j(1),\ldots,j(r)}$ as

$$[c, c, d].$$

If we apply (3.4) "at" the first pair of indices, we find that the range of the summation-term is vacuous, and we get

$$[2c, d]. \tag{3.10}$$

We could simplify this further, but let us instead turn our attention to the other way of applying (3.6), namely, to the last two indices of our original term. Here, as in Case 1, the short term vanishes because it still has an index c, while to the other terms we can make another application of (3.4) using the untouched index c:

$$\begin{aligned}[c, c, d] &= [c, c+d] - \sum_{c<\lambda\leqslant d} [c, \lambda, c+d-\lambda] \\ &= 0 - \sum_{c<\lambda\leqslant d} [c+\lambda, c+d-\lambda] \\ &\quad + \sum_{c<\mu\leqslant\lambda\leqslant d} [\mu, c+\lambda-\mu, c+d-\lambda].\end{aligned}$$

Let us reindex each of the above sums. In the first, let us take $\mu = c + \lambda$; in the second, $\nu = c + \lambda - \mu$. We also discard terms from the first summation in which the second distinguished index is $<2c$. Thus we get

$$- \sum_{2c < \mu \leqslant d} [\mu, 2c + d - \mu] + \sum_{c < \mu \leqslant d, c \leqslant \nu \leqslant c + d - \mu} [\mu, \nu, 2c + d - \mu - \nu]. \tag{3.11}$$

We now see that in the second sum, for each value of μ, the sum over ν reduces by (3.4) to $[\mu, 2c + d - \mu]$. Those terms with $\mu < 2c$ vanish by (3.2), so we are left with a summation over $2c \leqslant \mu \leqslant d$. Comparing with the first term of (3.11), we see that they cancel except for the $\mu = 2c$ term, leaving us with precisely (3.10).

This completes the verification that the procedure (3.2)–(3.4) is well defined. From this we get

LEMMA 6. *The construction* (3.2)–(3.4) *yields precisely all systems of elements* $a_{j(1),\dots,j(r)}$ $((j(1),\dots, j(r)) \in J_{n,m})$ *satisfying* (2.11). *Hence such a system is determined by arbitrarily prescribing those values* $a_{j(1),\dots,j(r)}$ *with-* $h(j(1),\dots, j(r)) = 0$.

Proof. Our construction introduces (3.6), i.e., the desired condition (2.11) as a definition wherever all subscripts of our element are $\geqslant 2^{n-r}$, while we have seen that those cases of (3.6) with any subscript $< 2^{n-r}$ reduce to $0 = 0$ in view of the other definitions. ∎

When $m \geqslant 2^{n-1}$, we have $h(m) = 0$ so we can get our system to satisfy $a_m = 1$, i.e., (2.12), as well. By Lemma 5, this yields

THEOREM 7. *Any k-algebra R satisfying* $R^{n+1} = \{0\}$ *can be embedded in a k-algebra* $H^{\oplus}$, *where H is a* $\mathbb{Z}$-*graded k-algebra zero in all degrees except* $1,\dots, 2^{n-1}$.

Hence by [1, *Theorem* 1], *R can also be embedded in strictly upper triangular* $(2^{n-1} + 1) \times (2^{n-1} + 1)$ *matrices over an associative k-algebra, which can even be taken commutative.* ∎

4. A SORT OF CONVERSE TO LEMMA 5

The above way of solving the system of equations (2.11) and (2.12) was suggested by the following observation, which is a sort of converse to Lemma 5, in that it shows that the existence of certain embeddings entails the existence of such systems $a_{j(1),\dots,j(r)}$.

LEMMA 8. *Let k be a commutative ring and n, m positive integers. Let*

$C = k[t]$ be the polynomial ring in one indeterminate over k, and let R be the C-algebra $tC/t^{n+1}C$.

Suppose H is a graded C-algebra zero in all degrees except $1,\ldots, m$, and $f\colon R \to H^{\oplus}$ is a k-algebra embedding such that $f(t^n)$ has nonzero component in H_m; and in fact that there exists a k-linear functional $\beta\colon H_m \to k$ sending this component to 1. (The existence of such a β clearly follows from the preceding condition if k is a field.)

Then if we write $f(t) = x = \sum x_i$, and define elements of k,

$$a_{j(1),\ldots,j(r)} = \beta(t^{n-r}x_{j(1)} \cdots x_{j(r)}), \tag{4.1}$$

these will satisfy (2.11) *and* (2.12).

Proof. We note that $tx = x^2$; hence $tx_i = \sum x_\mu x_{i-\mu}$. The assertions are now easily checked. ∎

The first idea this suggested was that one should look at the universal graded C-algebra U associated as in Section 2 above with the C-algebra $R = tC/t^{n+1}C$ of the above Lemma, find a normal form for its elements, say, by the method of the "Diamond Lemma" [3], and then write down explicitly the desired functional β. But the application of the method of [3] to that algebra U turned out to be particularly messy. (If we try to carry it out with C as base ring, the lack of a free basis or obvious C-module splitting makes for difficulty, while if we work over k, the centrality of t complicates our reduction system.) Nevertheless, consideration of what such a normal form *should* look like suggested how the $a_{j(1),\ldots,j(r)}$ might be defined, leading to (3.2)–(3.4). Further notes for readers familiar with [3]: It is clear that the calculations of the preceding section are "Diamond Lemma-type" arguments. Why, however, did we have a nontrivial computation to do in case 1 above, while in the analogous situations in [3] it is automatic that reductions done on non-overlapping parts of a monomial are independent? Because the form of the reductions we were doing here depended on the length r of the index-string in question, and the output of a reduction included terms with index-strings of different length, on which the available reductions were therefore not exactly the same as before. One may note the similarity of the set of $a_{j(1),\ldots,j(r)}$ we can prescribe arbitrarily by (3.3) to the set of basis-elements under the normal form obtained when k is a field in [2, Theorem 2]; but also the difference: the dependence of the allowed set of subscripts on the length of the string.

5. Application to Filtered Algebras

Now let R be a $\mathbb{Z}$-filtered k-algebra, in the sense of [2, Section 1]. Let T be the graded k-algebra generated by infinitely many k-module copies of R, one in each integer degree

$$\theta_i(R) \subseteq T_i \qquad (i \in \mathbb{Z}). \tag{5.1}$$

Let U now denote the quotient of T by the ideal generated by two families of elements; first, the union of the sets

$$\theta_i(R_{(i+1)}) \qquad (i \in \mathbb{Z}), \tag{5.2}$$

and secondly the elements

$$\theta_i(xy) - \sum_{v(x) \leqslant \mu \leqslant i - v(y)} \theta_\mu(x)\, \theta_{i-\mu}(y). \tag{5.3}$$

Note that in view of (5.2), the range of summation in (5.3) can be extended to any larger set of values without affecting the force of the condition; e.g., if $v(x)$ and $v(y)$ are positive it can be extended to $0 < \mu < i$, as in (2.5).

If we let $\varphi_i(x)$ denote the image of $\theta_i(x)$ in U_i, then the map $\varphi: R \to U^{\hat{\oplus}}$ given by

$$\varphi(x) = \sum_i \varphi_i(x) \tag{5.4}$$

is seen to be universal among all filtered algebra homomorphisms $R \to H^{\hat{\oplus}}$ (H a graded k-algebra).

The example of Corollary 4 (Section 1) shows that in trying to construct maps of filtered algebras R into algebras $H^{\hat{\oplus}}$, it may be desirable to assume that all elements of R have positive degree, i.e.,

$$R = R_{(1)}. \tag{5.5}$$

Under this hypothesis we can simplify the construction of U by first limiting the generating k-modules (5.1) to those with $i > 0$ since the other components are killed off when we divide by (5.2) for $i \leqslant 0$, and then restricting the relations (5.2) in the same way. With this modification, a homogeneous components T_m is the sum of the finite number of tensor product modules given earlier in (2.3) (with $i = m$); and U_m is the quotient of this sum by the span of the elements given earlier as (2.6), together with the sets

$$\theta_{i(1)}(R) \otimes \cdots \otimes \theta_{i(p)}(R_{(i(p)+1)}) \otimes \cdots \otimes \theta_{i(q)}(R) \qquad (i(1) + \cdots + i(q) = m). \tag{5.6}$$

Now suppose n and m are positive integers for which we have a system of elements $a_{j(1),\ldots,j(r)} \in k$ $((j(1),\ldots, j(r)) \in J_{n,m})$ satisfying (2.11) and (2.12), and that we define $\alpha: T_m \to R$ as in (2.7). By (2.12) α will satisfy

$\alpha \circ \theta_m = id_R$, and by (2.11) will annihilate all the relators (2.6) with $q < n$; α also clearly annihilates such relators with $q > n$, since it involves no $\pi_{j(1),\ldots,j(r)}$ acting on a component of T_m of length $r > n$. Now α will *not*, in general, annihilate relators (2.6) with $q = n$, nor all the relator-sets (5.6). But we claim it will at least take these into $R_{(n+1)}$. First, we observe that the image of a relator (2.6) with $q = n$ under any map $\pi_{j(1),\ldots,j(r)}$ will lie in $R^{q+1} = R^{n+1} = R_{(1)}^{n+1} \subseteq R_{(n+1)}$. As for a relator-set (5.6), α can be nonzero on it only if $\pi_{i(1),\ldots,i(r)}$ has nonzero coefficient in α. By (2.15) this entails

$$i(p) \geqslant 2^{n-r} \text{ which we note is } \geqslant n - r + 1. \tag{5.7}$$

Hence the image of (5.6) will lie in

$$R \cdots RR_{(i(p)+1)}R \cdots R \subseteq R \cdots RR_{(n-r+2)}R \cdots R, \tag{5.8}$$

with $r - 1$ "R's" and one $R_{(n-r+2)}$. Writing each R as $R_{(1)}$, we see that the product again lies in $R_{(n+1)}$ as claimed.

It follows that though $\alpha\colon T_m \to R$ will not in general factor through U_m, the composed map $T_m \to^{\alpha} R \to R/R_{(n+1)}$ will, giving a map $\bar{\alpha}\colon U_m \to R/R_{(n+1)}$ such that the composition $R \to^{\varphi_m} U_m \to^{\bar{\alpha}} R/R_{(n+1)}$ is just the quotient-map. Hence the kernel of φ_m is contained in $R_{(n+1)}$. This means that given $x \in R$ with $v(x) \leqslant n$, so that $x \notin R_{(n+1)}$, we have $\varphi_m(x) \neq 0$, hence $v(\varphi(x)) \leqslant m$.

Since given n we know that a system of elements $a_{j(1),\ldots,j(r)}$ satisfying (2.11) and (2.12) exists for $m = 2^{n-1}$, we can conclude:

THEOREM 9. *Let R be filtered k-algebra satisfying $R = R_{(1)}$, and let $\varphi\colon R \to U^{\hat{\oplus}}$ be the universal homomorphism of filtered k-algebras constructed above. Then for all $x \in R$,*

$$v(x) \leqslant v(\varphi(x)) \leqslant 2^{v(x)-1}. \quad \blacksquare \tag{5.9}$$

6. Some Minor Improvements

Note that we used a weak estimate on 2^{n-r} in the above argument (at (5.7)). Although we know by the examples in Section 1 that we cannot reduce the estimate $2^{v(x)-1}$ in the Theorem, we can take this slack out of our argument by a change on the other end:

Let us consider the universal homomorphism f of R into an algebra $H^{\hat{\oplus}}$ such that $v(f(x)) \geqslant 2^{v(x)-1}$. This means replacing the relator-sets (5.2) by the system of larger sets

$$\theta_i(R_{([\log_2 i]+2)}) \qquad \text{(where [ ] denotes "greatest integer } \leqslant\text{")}. \tag{6.1}$$

Now taking α as before, for any $\pi_{i(1),\ldots,i(r)}$ having nonzero coefficient in α, observe that each relator-set based on (6.1) lying in the tensor-summand on which this operator acts is carried by it to $R_{\langle 1\rangle}R_{\langle 1\rangle}\cdots R_{\langle [\log_2 i(p)]+2\rangle}\cdots R_{\langle 1\rangle}\subseteq R_{\langle [\log_2 i(p)]+r+1\rangle}$, which is contained in $R_{\langle n+1\rangle}$ by the first inequality of (5.7).

Thus we again get $v(f(x))\leqslant 2^{v(x)-1}$. But by construction we have made $v(f(x))\geqslant 2^{v(x)-1}$, so $v(f(x))=2^{v(x)-1}$ exactly.

Note also that in these proofs we have not used the full strength of R being filtered. The only cases of the condition $R_{\langle i\rangle}R_{\langle j\rangle}\subseteq R_{\langle i+j\rangle}$ that we have used are where i or j is 1.

It follows that if we start with a k-algebra R with a chain of subsets $R=R_{\langle 1\rangle}\supseteq R_{\langle 2\rangle}\supseteq\cdots$ satisfying

$$\text{each } R_{\langle i\rangle} \text{ is a } k\text{-submodule of } R, \tag{6.2}$$

$$RR_{\langle i\rangle}+R_{\langle i\rangle}R\subseteq R_{\langle i+1\rangle}, \tag{6.3}$$

$$\bigcap_i R_{\langle i\rangle}=\{0\}, \tag{6.4}$$

and define

$$w(x)=\sup\{i\mid x\in R_{\langle i\rangle}\}, \tag{6.5}$$

then we can get a homomorphism of R into a k-algebra $H^{\hat{\oplus}}$ such that $v(f(x))=2^{w(x)-1}$. Systems satisfying (6.2)–(6.4) arise naturally if we look at positive integer valued filtration functions v that do not assume all values. If v assumes only the values $b(1)<b(2)<\cdots$, and we define

$$R_{\langle i\rangle}=R_{(b(i))}, \tag{6.6}$$

then (6.2)–(6.4) clearly hold. The application of the above arguments to this system (6.6) gives a homomorphism f with $v(f(x))=2^{w(x)-1}$, which is generally smaller than $2^{v(x)-1}$. Assuming $b(i)\leqslant 2^{i-1}$ this will be a homomorphism of filtered algebras with respect to the original filtration given on R.

On the other hand, we may not always want the embedding that gives elements the smallest order; we may want to control the orders in some other way. To free ourselves from the restriction to powers of 2, we need to generalize the construction of Section 3.

COROLLARY 10 (to proof of Lemma 6). *Let $c(0)$, $c(1)$,... be positive integers such that for all i,*

$$c(i+1)\geqslant 2c(i), \tag{6.7}$$

and let k be any commutative ring. Then for any $n\geqslant 1$ and $m\geqslant c(n-1)$, *there exist elements* $a_{j(1),\ldots,j(r)}$ $((j(1),\ldots,j(r))\in J_{n,m})$ *satisfying* (2.11), (2.12), *and also the condition*

$$\textit{If } a_{j(1),\ldots,j(r)}\neq 0 \textit{ then } j(1),\ldots,j(r)\geqslant c(n-r) \text{ (cf. (2.15))}. \tag{6.8}$$

Notes on proof. We mimic the proof of Lemma 6, generally replacing 2^i by $c(i)$, in particular in (3.1) where the induction-index h is defined, and in (3.4), equivalently (3.6), which becomes

$$[c(n-r),d]=[c(n-r)+d]-\sum_{c(n-r)<\mu\leqslant d}[\mu,c(n-r)+d-\mu]. \tag{6.9}$$

When we follow through the proof, we see that most of the steps whose translations require some relation between successive values $c(n-r)$ and $c(n-r+1)$ are in the arguments saying that we can drop certain terms because they involve $r-1$ indices at least one of which is $<2c$. Clearly if for $<2c$ we read $<2c(n-r)\leqslant c(n-r+1)$, we can still drop those terms. In the one situation where we apply (6.9) with a "contracted" string of indices, namely, in going from (3.8) to (3.9), we see that if we simply cut the range of summation down by $c(n-r+1)-2c(n-r)$ on each side, this still works. (Actually, that step only uses $c(i+1)>c(i)$.) ∎

We now combine this with the two other improvements on Theorem 9 indicated earlier, in

PROPOSITION 11. *Let R be a k-algebra with a chain $R=R_{(1)}\supseteq R_{(2)}\supseteq\cdots$ satisfying* (6.2)–(6.4), *and let $w\colon R\to\{1,2,\ldots;+\infty\}$ be the function defined by* (6.5). *Let $c(0), c(1),\ldots$ be a sequence of integers satisfying* (6.7), *and define* $c(+\infty)=+\infty$.

Then there exists a graded k-algebra H and a k-algebra homomorphism $f\colon R\to \hat{H}^{\oplus}$ satisfying

$$v(f(x))=c(w(x)-1). \quad ∎ \tag{6.10}$$

This result includes Theorem 9 and the intermediate results we noted. Indeed:

COROLLARY 12. *Let R be a filtered ring satisfying $R=R_{(1)}$. Then:*

(i) *If the values of the filtration function v on $R-\{0\}$ lie in a family of positive integers $b(0)<b(1)<\cdots$, and if $c(0)<c(1)<\cdots$ are positive integers satisfying* (6.7), *then there exists a graded k-algebra H and a k-algebra homomorphism $f\colon R\to \hat{H}^{\oplus}$ such that when $v(x)=b(i)$, one has $v(f(x))=c(i)$. Thus, if for all i, $c(i)\geqslant b(i)$, this f is a homomorphism of filtered algebras.*

As special cases we have:

(ii) *If the values of* v *lie in a family* $c(0) < c(1) < \cdots$ *satisfying* (6.7), *then* f *can be taken to satisfy* $v(f(x)) = v(x)$.

(iii) *Without assumption on the values of* v, f *can be taken to satisfy* $v(f(x)) = 2^{v(x)-1}$. ∎

There are still a few tricks one can play in this line. For instance, if we are given a filtration function v with values in $\{2, 3, 4, \ldots\}$, then $[(v(x)+1)/2]$ is again a positive integer valued filtration function (cf. [2, Section 2.4]) so we can get $f: R \to \hat{H}^{\oplus}$ such that $v(f(x)) = 2^{[(v(x)+1)/2]}$. Note that $2^{[(v(x)+1)/2]} \geqslant v(x)$, through this is not true of $[(v(x)+1)/2]$ itself. The universal filtered algebra homomorphism $R \to U^{\oplus}$ will therefore satisfy $v(x) \leqslant 2^{[(v(x)+1)/2]}$, which is better than we could have gotten directly from any of the above results.

7. Some Observations and Questions

One may note the occurrence of the function 2^{n-1} first in our counterexamples (Section 1) and then in our positive results, and ask whether the base 2 is really so special, or whether there are variants of our problem that give, say, powers of 3 as best estimates.

We have some partial answers. In Lemma 1, if we replace (1.1) by

$$x^3 = tx \tag{7.1}$$

we get the estimate

$$v(x^{2n+1}) \geqslant 3^n. \tag{7.2}$$

The natural example of an algebra in which we have (7.1) is $R = sk[s]$, considered as an algebra over $C = k[s^2]$, with $x = s \in R$ and $t = s^2 \in C$. The construction analogous to that of Lemma 8 for this ring yields systems $a_{j(1),\ldots,j(r)}$ satisfying

$$a_{i(1),\ldots,i(p),\ldots,i(q)} = \sum_{\lambda,\mu} a_{i(1),\ldots,\lambda,\mu,i(p)-\lambda-\mu,\ldots,i(q)}. \tag{7.3}$$

These in turn correspond via considerations like those of Section 2 to k-linear maps $R \to \hat{H}^{\oplus}$ which, rather than satisfying the multiplicative condition for a ring homomorphism

$$f(xy) = f(x)\,f(y), \tag{7.4}$$

satisfy the weaker condition

$$f(xyz) = f(x)\,f(y)f(z). \tag{7.5}$$

And indeed, the study of embeddings satisfying (7.5) or the analogous condition with some larger integer in place of "3" involves estimates of the sort (7.2). We don't know of any great per se interest in maps satisfying conditions like (7.5), but in fact, after the above passage was written, these observations turned out to be of heuristic value in discovering the main result of [4] (see motivation sketched in [4, Section 3]), though they do not appear in the final version of the proof [4, Section 4].

Finally, let us look again at our commutative base-ring k. For simplicity, we limit attention to the case where

$$k \text{ has no idempotent elements other than 0 and 1.} \tag{7.6}$$

When k is a field, we know from [2] that we can get graded embeddings with $v(f(x)) = v(x)$. On the other hand the examples of Section 1 are easily strengthened (cf. (1.9)) to show that assuming (7.6), the bound $v(f(x)) \leqslant 2^{n-1}$ is best if k either has non-nilpotent nonunits, or has non-nilpotent nil ideal. This leaves only the case where k is a local ring with nilpotent maximal ideal,

$$\mathfrak{m}^{N+1} = \{0\} \qquad (N > 1, \mathfrak{m}^N \neq \{0\}). \tag{7.7}$$

Examples are $k = \mathbb{Z}/p^{N+1}\mathbb{Z}$ (p a prime) and $k = k_0[t]/t^{N+1}k_0[t]$ (k_0 a field). It would be interesting to know what kind of best estimates of $v(f(x))$ hold for such base rings: exponential growth, as for "most" k; linear growth, as for k a field; or something in between, such as polynomial growth.

Acknowledgments

This work was done while authors Britten and Lemire were partially supported by NRC Grants A8471 and A7742, and Bergman by NSF Contract MCS 80-02317. The authors wish to acknowledge the help of Mr. T. Bass, who wrote a computer program for authors Britten and Lemire to answer some questions arising at an early stage in this investigation.

References

1. G. M. Bergman, Embedding rings in completed graded rings. 1. Triangular embeddings, *J. Algebra* **84** (1983), 14–24.
2. G. M. Bergman and S. M. Vovsi, Embedding rings in completed graded rings. 2. Algebras over a field, *J. Algebra* **84** (1983), 25–41.

3. G. M. BERGMAN, The Diamond Lemma for ring theory, *Adv. in Math.* **29** (1978), 178–218.
4. G. M. BERGMAN, Embedding rings in completed graded rings. 4. Commutative algebras, *J. Algebra* **84** (1983), 62–106.
5. J. E. HUMPHREYS, "Introduction to Lie Algebras and Representation Theory," Graduate Texts in Mathematics, 9, Springer-Verlag, New York/Berlin, 1972.
6. R. L. KRUSE AND D. T. PRICE, "Nilpotent Rings," Gordon & Breach, New York, 1969.